UDC 621.395.97:
621.395.452

BBC

1969
THE QUEEN'S AWARD
TO INDUSTRY

RESEARCH DEPARTMENT

REPORT

The subjective effect of frequency shift in the transmission of high-quality sound signals

No. 1969/13

Research Department, Engineering Division
THE BRITISH BROADCASTING CORPORATION

CORRIGENDUM

RESEARCH DEPARTMENT – BRITISH BROADCASTING CORPORATION

Research Report No. 1969/13

THE SUBJECTIVE EFFECT OF FREQUENCY SHIFT IN THE TRANSMISSION OF HIGH-QUALITY SOUND SIGNALS

Please add to title page:

Work covered by this report was undertaken by the BBC Research Department for the BBC and the ITA

GDM
24.10.69

RESEARCH DEPARTMENT

THE SUBJECTIVE EFFECT OF FREQUENCY SHIFT IN THE TRANSMISSION OF HIGH-QUALITY SOUND SIGNALS

Research Department Report No. 1969/13
UDC 621.395.97:
621.395.452

It uses SI units in accordance with B.S. document PD 5686.

W.I. Manson, B.Sc. (Eng.)

D. Maurice

Head of Research Department

(EL-26)

Research Department Report No. 1969/13

THE SUBJECTIVE EFFECT OF FREQUENCY SHIFT IN THE TRANSMISSION OF HIGH-QUALITY SOUND SIGNALS

July 1969

Research Department Report No. 1969/13
UDC 621.395.97:
621.395.452

THE SUBJECTIVE EFFECT OF FREQUENCY SHIFT IN THE TRANSMISSION OF HIGH-QUALITY SOUND SIGNALS

SUMMARY

This report describes tests carried out to determine the subjective effect, on the reproduction of high-quality sound signals, of shifting the frequency of each spectral component by the same amount.

The results obtained are discussed and suggestions made for a working tolerance.

1. INTRODUCTION

The means now used to distribute sound programme signals include systems using suppressed-carrier single-sideband transmission. At the receiving terminal a local oscillator is used to restore the carrier suppressed at the sending terminal. The frequency of the restored carrier is nominally the same as that of the suppressed carrier; slight discrepancies do occur, however, and, as a result, the received baseband spectrum is subjected to a frequency shift equal to the difference between the two carriers.

This effect differs from that introduced when a recording is reproduced at incorrect speed. The change in frequency, in the case of incorrect replay speed, is multiplicative – all signal components are scaled in frequency by a constant factor and harmonic relationships are undisturbed. In the case of carrier frequency error, on the other hand, the effect is additive – the frequency of each component of the baseband signal is raised, or lowered, by the same amount, and consequently the precise relationship between signal components in a harmonic series is destroyed. The subjective effect of this additive frequency shift on transmitted music and speech signals is the introduction of various unusual forms of distortion – the particular form of the distortion depending on the nature of the programme material being transmitted.

For high-quality music circuits a tolerance limit of ±2 Hz shift is at present generally accepted, but no evidence to justify this figure has been found in the literature. Tests have therefore been carried out to determine the subjective effects of various amounts of frequency shift and thus to provide a basis for tolerances to be specified for high-quality transmission systems. This report discusses the tests carried out and the results obtained, and suggests a working tolerance.

2. DESCRIPTION OF EXPERIMENT

2.1. Selection of Test Material

Before starting formal tests a preliminary investigation was carried out to select suitable programme items and, further, to discover any factors which might influence the choice of test procedure. [See Appendix I]

For the purpose of this investigation a 12 kHz channel, consisting of modulating and demodulating equipment, was made available by the G.P.O. One of the fixed-frequency carrier oscillators in this equipment was replaced by a stable variable-frequency oscillator, which was adjustable over a range of ± 17 Hz about the nominal figure.

A wide range of programme material was reproduced through the carrier channel with various degrees of frequency shift applied. As the amount of shift was increased from zero, the first effect to be detected when the original and frequency-shifted signals were compared* was a pitch change. With a further increase in frequency shift other forms of impairment became apparent even without comparison with the original programme. Impairment was found to be more noticeable on programme material in which low frequencies rather than high frequencies were predominent; this result is consistent with the greater percentage change of frequency – and hence the greater change of pitch – suffered by low-frequency signals for a given shift in the baseband spectrum. The precise nature of the impairment depended on the programme item and on the amount of frequency shift applied – in some cases tone quality was altered, while in others chords were rendered inharmonic, the relative pitch of notes in a scale was disturbed or beating effects were introduced. Even in the most extreme cases, however, all the effects heard were confined to the low and middle frequency range; it was therefore concluded that the restriction of the channel bandwidth to 12 kHz was unlikely to influence the results.

The final choice of items for the subjective tests was made with the aid of two skilled observers, whose day-to-day duties require a professional standard of musical appreciation. Four test passages – a string quartet excerpt with a prominent cello part, a scale played on a double bass, an organ excerpt and a scale in double octaves played on a piano – were selected as representative of the most critical material. Each test passage lasted between 18 and 42 seconds.

* For reasons discussed in Appendix I direct comparison of impaired and unimpaired signals was not allowed in the final tests.

As a guide to the range of impairment to be explored, it was noted that, with the more critical programme material, the effect was usually considered obvious with a 8 Hz shift, generally detectable with 4 Hz shift, and often imperceptible with a 2 Hz shift.

2.2. Observers

Two groups of observers took part in the tests. The first, which will be referred to as the 'amateur' group, comprised twelve members of Research Department staff, all of whom had a keen interest in music either as listener or as an amateur performer. The second, a 'professional' group, also twelve in number, consisted of BBC studio staff whose normal duties require a professional standard of musical appreciation.

2.3. Test Conditions

All the listening tests were carried out in quiet rooms having acoustics similar to those of an average domestic living room. Two to six observers attended each session and heard the test recordings reproduced on a BBC monitoring loudspeaker, type LS5/5, at a level of some 83* phons maximum.

2.4. Presentation of Tests

The test procedure adopted was largely determined by effects noted during the preliminary investigation; these are discussed in Appendix I. The programme of tests was divided into four sessions; two of the sessions were devoted to positive frequency shift and two to negative frequency shift.

Each test session consisted of 24 items, each of the four selected passages being presented with frequency shifts of 0, 1, 2, 4, 8 and 16 Hz. These items were presented in an arbitrary order but repetition of any one passage or any one degree of frequency shift in consecutive items was avoided.

In one of the sessions devoted to positive frequency shift and one devoted to negative frequency shift the order of the 24 items was reversed.

All observers made assessments of all four sessions.

Every test session began with a recorded introduction which included for illustration, a demonstration of the effect, on the test passages, of a large shift (17 Hz, positive or negative as appropriate); this was followed by a demonstration of the excerpts unimpaired for reference.

The observers were then asked to judge the impairment of the 24 test items and to allocate marks according to the following scale:—

Impairment	Grade
Imperceptible	1
Just perceptible	2
Definitely perceptible but not disturbing	3
Somewhat objectionable	4
Definitely objectionable	5
Unusable	6

The instructions to the observers are given in full in Appendix II.

* Measured on B and K sound level meter, type 2203, with weighting A.

The total duration of each test session, including the introduction, was about 20 minutes.

The tests carried out by Research Department staff were organised so that the observers attended only two sessions, spaced by about ½ hour, in any one day. Due to programme commitments it was impossible to arrange the tests in this way for the professional observers; here each observer completed all the tests in one day, attending two morning and two afternoon sessions.

Having decided the form of presentation, the necessary test recordings were prepared, again making use of the GPO carrier channel terminal equipment to give the various amounts of frequency shift required. The frequency shift introduced was checked both before and after recording by measuring the frequency of the output signal when an accurate 1 kHz reference tone was applied to the input, and was in every case within 0·1 Hz of its nominal value. The recordings prepared in this way were subsequently edited* and grouped into the four tapes for the four test sessions as described.

3. RESULTS

As stated in Section 2.4, each test excerpt was judged twice by every observer. The average of the two results thus obtained was taken as the effective grading for the observer and was used in subsequent analysis.

Table 1(*a*) and (*b*) shows for amateur and professional observers respectively the mean gradings, the standard deviations and the standard errors of the means for the various test conditions.

There are minor differences in the gradings for positive and negative frequency shift, but no marked systematic bias is evident in the results. Since the two conditions are equally probable it is appropriate to consider further only the less favourable grading of the two obtained; this course has been adopted in preparing Figs. 1 and 2 which show the mean gradings plotted as a function of frequency shift for the amateur and professional observers respectively. The standard errors of the means are shown to indicate the degree of confidence to be placed in the various points.

Interpretation of the curves in Figs. 1 and 2 is rendered difficult by the fact that, while some observers graded the impairment for the nominal zero frequency shift (which in fact was always less than ± 0·1 Hz) as 'Imperceptible', most of them found fault with the reproduction even in this condition. Such a result may be expected to arise from over-anxiety on the part of the observer or through failure to distinguish the impairment being investigated from other faults, musical or technical. The same factors would also affect the assessment of small amounts of impairment, and it is therefore not surprising that some of the mean gradings shown for shifts of 1 Hz — and in certain cases 2 Hz — also appear anomalous, though these points, as well as the grading for zero frequency shift, are statistically no less reliable than the rest. Examples of extraneous effects which could be present are imperfect intonation in the original performance and random frequency fluctuations introduced by recording and copying.

* Much of the preparation of the test tapes, the presentation of the tests and the analysis of results was carried out by G.R. Mitchell.

TABLE 1

Mean Grading (M), Standard Deviation (SD) and Standard Error of Mean (SE) for various amounts of Frequency Shift
(a) Amateur Observers – (b) Professional Observers

			0 Hz			1 Hz			2 Hz			4 Hz			8 Hz			16 Hz		
		Test Item	M	SD	SE	M	SD	SE	M	SD	SE	M	SD	SE	M	SD	SE	M	SD	SE
(a)	Frequency Increased	String Quartet	1·50	0·58	0·17	1·33	0·47	0·14	1·50	0·74	0·22	1·88	0·74	0·22	2·42	0·93	0·28	4·21	1·88	0·57
		Double Bass	2·00	0·79	0·24	1·63	0·51	0·15	2·17	1·03	0·31	2·83	0·90	0·27	4·33	0·85	0·26	5·75	0·60	0·18
		Organ	2·13	0·87	0·26	2·21	0·99	0·30	2·50	1·08	0·33	2·04	0·95	0·29	4·17	0·94	0·28	5·63	0·85	0·26
		Piano	1·58	0·57	0·17	1·92	0·76	0·23	1·46	0·59	0·18	1·83	0·66	0·20	4·33	1·25	0·38	5·50	1·19	0·36
	Frequency Decreased	String Quartet	1·63	0·42	0·13	1·58	0·76	0·23	1·63	0·65	0·20	1·54	0·48	0·14	2·08	0·76	0·23	3·46	1·65	0·50
		Double Bass	1·83	0·62	0·19	2·38	1·33	0·40	2·67	1·31	0·40	2·58	1·02	0·31	4·63	1·02	0·31	5·79	0·69	0·21
		Organ	1·88	0·82	0·25	2·04	0·95	0·29	1·92	0·86	0·26	2·58	1·10	0·33	4·46	1·31	0·40	5·46	0·85	0·26
		Piano	1·71	0·75	0·23	2·08	1·04	0·31	2·04	0·78	0·23	2·21	0·92	0·28	4·33	1·40	0·42	5·71	0·83	0·25
(b)	Frequency Increased	String Quartet	1·75	0·48	0·14	1·88	0·77	0·23	1·79	0·56	0·17	2·58	1·02	0·31	3·83	0·87	0·26	5·46	0·56	0·17
		Double Bass	2·21	0·59	0·18	2·21	0·72	0·22	2·25	0·99	0·30	3·04	0·83	0·25	5·08	0·79	0·24	6·00	0·00	0·00
		Organ	2·21	0·43	0·13	2·33	0·62	0·19	2·04	0·72	0·22	2·63	0·85	0·26	4·92	0·76	0·23	5·96	0·14	0·04
		Piano	1·79	0·85	0·26	2·33	0·69	0·21	2·54	1·05	0·32	2·58	1·00	0·30	5·04	0·63	0·19	5·96	0·14	0·04
	Frequency Decreased	String Quartet	2·25	0·80	0·24	2·17	0·69	0·21	2·21	0·66	0·20	2·67	0·85	0·26	3·46	0·88	0·26	5·46	0·85	0·26
		Double Bass	1·96	0·59	0·18	2·25	0·90	0·27	2·33	0·72	0·22	3·21	0·88	0·26	4·50	0·94	0·28	5·88	0·22	0·07
		Organ	2·33	0·83	0·25	2·67	1·09	0·33	2·25	0·66	0·20	3·17	1·09	0·33	4·67	0·83	0·25	5·83	0·31	0·09
		Piano	2·21	0·97	0·29	2·63	0·77	0·23	2·54	1·20	0·36	2·79	1·07	0·32	4·63	1·12	0·34	5·83	0·43	0·13

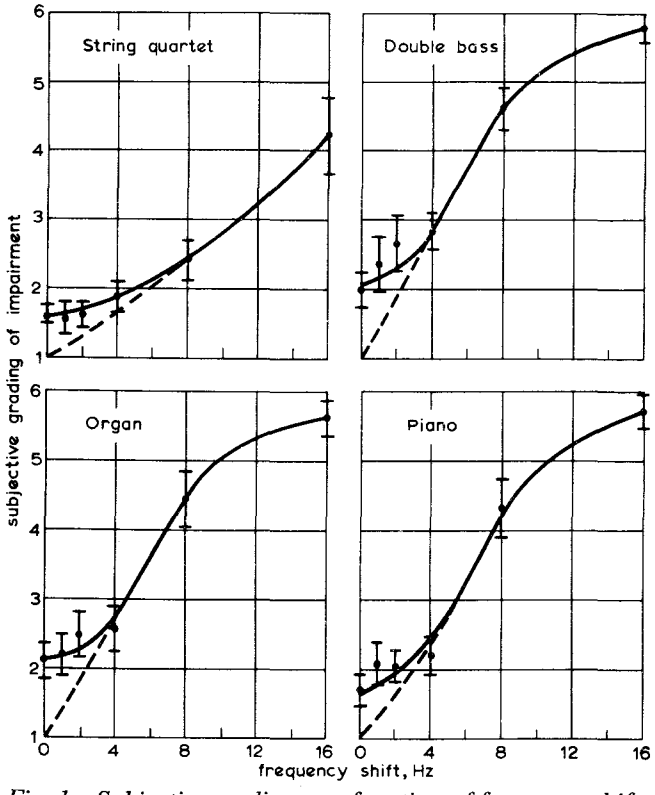

Fig. 1 - Subjective grading as a function of frequency shift: mean for amateur observers. 2 x standard error of mean shown thus ⟊ For each amount of frequency shift the result plotted is that obtained for the direction of shift giving the less favourable grading.

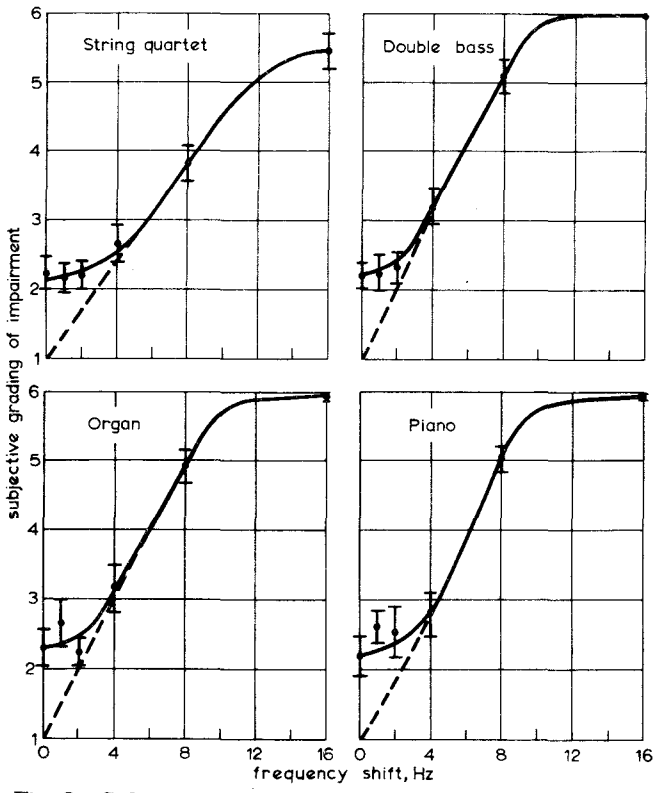

Fig. 2 - Subjective grading as a function of frequency shift: mean for professional observers. 2 x standard error of mean shown thus ⟊ For each amount of frequency shift the result plotted is that obtained for the direction of shift giving the less favourable grading.

The results which would have been obtained in the absence of such extraneous influences can, however, be estimated by redrawing each curve so as to pass through the point 0, 1 (zero subjective effect for zero frequency shift), ignoring, or giving less weight to, the gradings obtained for small amounts of frequency shift. The curves thus obtained are shown dotted in Figs. 1 and 2.

Particular interest is attached to the impairment produced by a frequency shift of 2 Hz since this is the upper limit of tolerance currently applied to high-quality circuits. The estimated mean gradings for 2 Hz frequency shift, taken from the dotted curves in Figs. 1 and 2 are given below for the various programme items.

TABLE 2

Assessments for 2 Hz Shift

	Amateur Observers		Professional Observers	
	Mean Grade	% above Grade 3½	Mean Grade	% above Grade 3½
String Quartet	1·4	0·3	1·7	0·3
Double Bass	1·9	11	2·0	6
Organ	1·9	7	2·0	2
Piano	1·7	1·2	1·8	8

It will be seen that for the two most critical items, double bass and organ, the mean grade for the professional observers is 2·0, while the amateur observers are slightly more tolerant, giving a mean grading of 1·9 for both items. The string quartet and piano items proved to be somewhat less stringent tests, yielding mean gradings of 1·4 and 1·7 respectively for the amateur observers and 1·7 and 1·8 for the professional observers.

Assuming a Gaussian frequency distribution[1] the values of standard deviation given in Table 1 for a 2 Hz frequency shift have been used with the mean gradings of Table 2 to estimate the percentage of observers whose grading falls in the 'unfavourable' region above grade 3½. The results are included in Table 2; it will be seen that the proportion of observers whose estimate falls in the region above 3½ is never more than 11%.

4. CONCLUSIONS

The effects introduced when the components of a high-quality sound signal are shifted in frequency have been investigated by subjective tests.

Interpretation of the data obtained for small amounts of frequency shift was complicated by various extraneous effects. Alternative test procedures, likely to reduce these undesired effects, can be envisaged but these would be much more time-consuming and costly than those adopted.

It can, however, be concluded that the ± 2 Hz limit of tolerance of frequency shift hitherto applied will give results generally acceptable to the great majority of listeners. For very critical programme material and test conditions, the proportion of observers who would rate the quality impairment due to a ± 2 Hz shift at more than grade 3½ on a 6-point scale is estimated to be only 11%. This result is regarded as acceptable and the present tolerance limit of ± 2 Hz shift is considered to represent good engineering practice.

5. ACKNOWLEDGEMENT

Acknowledgement is due to the GPO for making available the carrier channel equipment used to introduce frequency shift.

6. REFERENCE

1. A simple application to broadcasting of the results of subjective tests. BBC Research Department Report No. A-091, Serial No. 1966/43.

APPENDIX I

Test Procedure

The choice of a test procedure was complicated by the unusual nature of the impairments being studied. These impairments – poor intonation, unusual tonal quality, exaggerated beats and the like – all suggest peculiarities in the musical performance rather than defects in a transmission system. Faulty intonation in the original performance could in some cases be partially compensated by the frequency shift. Moreover the question of personal preference could conceivably arise; an observer might even prefer the impaired rendering of an excerpt, at least for small amounts of frequency shift and, if asked to apply grading on an absolute scale, would then place the impaired rendering above the unimparied. However, it is no part of the present investigation to study this aspect of the phenomenon of frequency shift, and the test procedure should therefore be designed to reduce the effects of personal likes and dislikes to a minimum.

The effects of individual preferences can be eliminated in subjective studies by determining the ability of the observer to detect a difference between the impaired and unimpaired versions of the test item presented in rapid succession. However, it was found during the preliminary investigation that a change of pitch was apparent in these circumstances even when the amount of frequency shift was such that no other impairment could be detected. The results of tests carried out in this way would be determined predominantly by the observers' ability to detect small differences in pitch between the two presentations. In normal listening, on the other hand, the unimpaired version is not available for reference, and in these circumstances most observers – excluding only those with a sense of absolute pitch – are much more tolerant of errors in frequency. It would, therefore, clearly be wrong to adopt a test procedure based on direct comparison.

For the reasons discussed above, test methods involving either grading of preference on an absolute scale, or close comparison of original and impaired programme items, are both undesirable. The procedure adopted was, in effect, a compromise between these two extremes. The unimpaired excerpts were demonstrated at the start of each test session, with an interval of ½ minute before the first test item, and the observers were instructed that any item indistinguishable from this original presentation should be graded '1'. Thus a measure of comparison was included in the tests and the effects of preference thereby reduced. However, because of the ½ minute time interval between the preliminary demonstration of the unimpaired excerpts and the start of the tests, there was little possibility of direct comparison.

APPENDIX II

Introductions to the Test Series

The introductions for the various sessions of subjective tests were prepared in recorded form and played, as appropriate, to the observers at the beginning of each session.

The introduction to the first session included the test excerpts, both in their original form, and after being subjected to gross frequency shift (17 Hz). The second test session was essentially similar to the first, the only difference being in the order of presentation of the test items; a shortened version of the introduction was therefore used and the examples given were restricted to a playing of the unimpaired excerpts for reference.

A full introduction was however included at the beginning of test session three, while session four, like session two, included only the shortened introduction.

The texts of the recorded introduction to sessions one, two and three are given below; the introduction to session four is identical with that to session two – apart from the session number.

Recorded introduction to the first session of Tests

'Certain kinds of transmission system are capable of producing an unusual form of distortion. The object of these subjective tests is to determine how much of this distortion can be tolerated.

Before starting the tests we will demonstrate the nature of the distortion by playing a number of programme excerpts, first in their original form, and then with the distortion present to a much greater degree than would ever occur in practice.

There are four items – String Quartet, Solo Double Bass, Organ and Piano.

First, the four items in their original form:–

(Demonstration)

'Now the same four items with the unusual form of distor-

tion referred to earlier. The amount of distortion present in this case is much greater than would ever occur in practice'.

(Demonstration)

'We now come to the test proper; first you will hear the same four excerpts again in their original form, for reference. Then, after a short interval, while the tape is changed, you will hear the same four items presented a number of times in random order and with various amounts of distortion. We want you to grade each item according to the amount of impairment of the type you have already heard demonstrated. The grading scale is given on the score sheet; any item which is indistinguishable from the original will, of course, be graded '1'.

Here are the four original items again for reference'.

(Demonstration)

Recorded introduction to the second session of Tests

'This is the second session of tests. The nature of the distortion being investigated in this session is the same as in the last, so it is not necessary to repeat the preliminary demonstration.

However, to refresh your memory, you will now hear the four items again in their original form. Then, after a pause while the tape is changed, you will hear another series of tests and we would like you, as before, to grade the items'.

(Demonstration)

Recorded introduction to the third session of Tests

'This is the third session of tests. In this session we are concerned with a form of distortion slightly different from that of the previous sessions so it is necessary to have another preliminary demonstration. You will now hear the four items used in the earlier tests; first in their original form:—'

(Demonstration followed by remainder of introduction as for session one)

SMW

Printed by BBC Research Department, Kingswood Warren, Tadworth, Surrey